Journal of Global Optimization **24:** 163–185, 2002.

Solving Standard Quadratic Optimization Problems via Linear, Semidefinite and Copositive Programming

IMMANUEL M. BOMZE[1] and ETIENNE DE KLERK[2]
[1]*ISDS, Universität Wien, Vienna, Austria; E-mail: immanuel.bomze@univie.ac.at;* [2]*Faculty of Information Technology and Systems, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands; E-mail: E.deKlerk@ITS.TUDelft.NL*

Abstract. The problem of minimizing a (non-convex) quadratic function over the simplex (the standard quadratic optimization problem) has an exact convex reformulation as a copositive programming problem. In this paper we show how to approximate the optimal solution by approximating the cone of copositive matrices via systems of linear inequalities, and, more refined, linear matrix inequalities (LMI's). In particular, we show that our approach leads to a polynomial-time approximation scheme for the standard quadratic optimzation problem. This is an improvement on the previous complexity result by Nesterov who showed that a 2/3-approximation is always possible. Numerical examples from various applications are provided to illustrate our approach.

Key words: Approximation algorithms; Stability number; Semidefinite programming; Copositive cone; Standard quadratic optimization; Linear matrix inequalities

1. Introduction

A standard quadratic optimization problem (standard QP) consists of finding global minimizers of a quadratic form over the standard simplex, i.e., we consider global optimization problems of the form

$$p^* := \min_{\mathbf{x}\in\Delta} \mathbf{x}^\top Q\mathbf{x} \tag{1}$$

where Q is an arbitrary symmetric $n \times n$ matrix; a $\top$ denotes transposition; and Δ is the standard simplex in the n-dimensional Euclidean $\mathbb{R}^n$,

$$\Delta = \{\mathbf{x} \in \mathbb{R}^n_+ : \mathbf{e}^\top \mathbf{x} = 1\}\,,$$

where $\mathbf{e} = [1, \dots, 1]^\top \in \mathbb{R}^n$ and $\mathbb{R}^n_+$ denotes the non-negative orthant in $\mathbb{R}^n$. To avoid trivial cases, we assume throughout the paper that the objective is not constant over Δ, which means that $\{Q, E_n\}$ are linearly independent where $E_n = \mathbf{e}\mathbf{e}^\top$ is the $n \times n$ matrix consisting entirely of unit entries, so that $\mathbf{x}^\top E_n \mathbf{x} = (\mathbf{e}^\top \mathbf{x})^2 = 1$ on Δ. We need some further notation: as usual, I_n denotes the $n \times n$ identity matrix, and $\mathbf{e}_i \in \mathbb{R}^n$ its ith column (i.e., the ith standard basis vector in $\mathbb{R}^n$).

For a review on standard QPs and its applications, see [3]. We only mention here that this problem is known to be NP-hard, and contains the max-clique problem in graphs as a special case. Note that the minimizers of (1) remain the same if Q is replaced with $Q + \gamma E_n$ where γ is an arbitrary constant. So without loss of generality assume henceforth that all entries of Q are non-negative. Furthermore, the question of finding minimizers of a general quadratic function $\mathbf{x}^\top A\mathbf{x} + 2\mathbf{c}^\top\mathbf{x}$ over Δ can be homogenized considering the rank-two update $Q = A + \mathbf{e}\mathbf{c}^\top + \mathbf{c}\mathbf{e}^\top$ in (1) which has the same objective values on Δ.

In this paper we will show how to derive approximation guarantees for this problem via semidefinite programming (SDP). The main idea is as follows: we can give an exact reformulation of the standard quadratic optimization problem as a copositive programming problem, and subsequently approximate the copositive cone using either linear inequality systems, yielding LP relaxations; or, more refined, systems of linear matrix inequalities (LMIs), yielding an SDP formulation. This methodology is due to De Klerk and Pasechnik [5] and Parrilo [14] (see also [15]). We will show that we obtain a polynomial-time ϵ-approximation for problem (1) for each $\epsilon > 0$ in this way. Such an approximation is known as a *polynomial-time approximation scheme* (PTAS). This improves on a result by Nesterov [10], who showed that a 2/3-approximation is always possible.

Both SDP and copositive programming problems are examples of conic programming problems, and we begin by reviewing these concepts.

1.1. PRELIMINARIES; CONIC PROGRAMMING

We define the following convex cones:

- The $n \times n$ symmetric matrices:
  $\mathcal{S}_n = \{X \in \mathbb{R}^{n\times n}, X = X^\top\}$;
- the $n \times n$ symmetric positive semidefinite matrices:
  $\mathcal{S}_n^+ = \{X \in \mathcal{S}_n, \mathbf{y}^\top X\mathbf{y} \geqslant 0 \quad \text{for all } \mathbf{y} \in \mathbb{R}^n\}$;
- the $n \times n$ symmetric copositive matrices:
  $\mathcal{C}_n = \{X \in \mathcal{S}_n, \mathbf{y}^\top X\mathbf{y} \geqslant 0 \quad \text{for all } \mathbf{y} \in \mathbb{R}_+^n\}$;
- the $n \times n$ symmetric completely positive matrices:
  $\mathcal{C}_n^* = \{x = \Sigma_{i=1}^k \mathbf{y}_i\mathbf{y}_i^\top, \mathbf{y}_i \in \mathbb{R}_+^n \ (i = 1, \dots, k)\}$;
- the $n \times n$ symmetrical nonnegative matrices:
  $\mathcal{N}_n = \{X \in \mathcal{S}_n, X_{ij} \geqslant 0 \ (i, j = 1, \dots, n)\}$;
- the $n \times n$ symmetric doubly nonnegative matrices:
  $\mathcal{D}_n = \mathcal{S}_n^+ \cap \mathcal{N}_n$.

We consider the usual inner product $\langle X, Y\rangle := \mathrm{Tr}(XY)$ on $\mathcal{S}_n$ and recall that the completely positive cone is the dual of the copositive cone, and that the nonnegative and semidefinite cones are seld-dual with respect to this inner product. Furthermore, the dual cone of $\mathcal{D}_n$ is $\mathcal{D}_n^* = \mathcal{S}_n^+ + \mathcal{N}_n$, a cone which is contained in $\mathcal{C}_n$ and which will play an important role in the relaxations to follow.

For a given cone $\mathcal{K}$ and its dual cone $\mathcal{K}^*$ we define the primal and dual pair of conic linear programs:

$$\begin{aligned}&(P)\ \ p^* := \inf_X\{\langle C, X\rangle : \langle A_i, X\rangle = b_i \ (i=1,\dots,m),\ X \in \mathcal{K}\}\\&(D)\ \ d^* := \sup_{\mathbf{y}}\left\{\mathbf{b}^\top\mathbf{y} : C - \sum_{i=1}^m y_i A_i \in \mathcal{K}^*,\ \mathbf{y}\in\mathbb{R}^m\right\}.\end{aligned} \tag{2}$$

If $\mathcal{K} = \mathcal{S}_n^+$ we refer to *semidefinite programming*, if $\mathcal{K} = \mathcal{N}_n$ to linear programming, and if $\mathcal{K} = \mathcal{C}_n$ to *copositive programming*.

The well-known conic duality theorem, see, e.g., [18], gives the duality relations between (P) and (D).

THEOREM 1.1 (Conic duality theorem). *If there exists an interior feasible solution $X^0 \in int(\mathcal{K})$ of (P), and a feasible solution of (D) then $p^* = d^*$ and the supremum in (D) is attained. Similarly, if there exist $\mathbf{y}^0 \in \mathbb{R}^m$ with $C - \sum_{i=1}^m y_i^0 A_i \in int(\mathcal{K}^*)$ and a feasible solution of (P), then $p^* = d^*$ and the infimum in (P) is attained.*

As is well known [11], optimization over the cones $\mathcal{S}_n^+$ and $\mathcal{N}_n$ can be done in polynomial-time (to compute an ϵ-optimal solution), but copositive programming is NP-hard, as we will see in the next section.

1.2. STANDARD QUADRATIC OPTIMIZATION VIA COPOSITIVE PROGRAMMING

In [4] it is shown that we can reformulate problem (1) as the copositive programming problem

$$p^* := \min\{\langle Q, X\rangle : \langle E_n, X\rangle = 1, X \in \mathcal{C}_n^*\}. \tag{3}$$

Problem (3) is called a copositive program because of its dual formulation

$$p^* := \max\{\lambda : Q - \lambda E_n \in \mathcal{C}_n,\ \lambda \in \mathbb{R}\} \tag{4}$$

(note that the optimal values of both (3) and (4) are attained and equal to Theorem 1.1; see [4]). The reformulation makes it clear that copositive programming is not tractable (see, e.g., [19, 4]). In fact, even the problem of determining whether a matrix is not copositive is NP-complete [9].

In [4], some ideas from interior point methods for semidefinite programming are adapted for the copositive programming case, but convergence cannot be proved. The absence of a computable self-concordant barrier for this cone basically precludes the application of interior point methods to copositive programming.

A solution to this problem was recently proposed by Parrilo [14], who showed that one can approximate the copositive cone to a given accuracy by a sufficiently large set of linear matrix inequalities. In other words, each copositive programming problem can be approximated to a given accuracy by a sufficiently large SDP. Of course, the size of the SDP can be exponential in the size of the copositive program.

In the next section we will review the approach of Parrilo, and subsequently work out the implications for the copositive formulation of the general quadratic optimization problem by applying the approach of De Klerk and Pasechnik [5]. The basic idea is to replace the copositive cone in (4) by an approximation: either a polyhedral cone or a cone defined by linear matrix inequalities. In this way we obtain a tractable approximation problem.

2. Approximations of the copositive cone

Since any $\mathbf{y} \in \mathbb{R}^n_+$ can be written as $\mathbf{y} = \mathbf{x} \circ \mathbf{x}$ for some $\mathbf{x} \in \mathbb{R}^n$ where $\circ$ indicates the componentwise (Hadamard) product, we can represent the copositivity requirement for an $(n \times n)$ symmetric matrix M as

$$P(\mathbf{x}) := (\mathbf{x} \circ \mathbf{x})^\top M (\mathbf{x} \circ \mathbf{x}) = \sum_{i,j=1}^{n} M_{ij} x_i^2 x_j^2 \geqslant 0 \quad \text{for all } \mathbf{x} \in \mathbb{R}^n . \tag{5}$$

There are many possible representations of the polynomial P as a homogeneous polynomial of degree four, if we allow for nonzero coefficients of terms like $(x_i x_j)(x_k x_l)$ for $i \neq j \neq k \neq l$.

In particular, if we represent $P(\mathbf{x})$ via

$$P(\mathbf{x}) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}} \tag{6}$$

where $\tilde{\mathbf{x}} = [x_1^2, \ldots, x_n^2, x_1 x_2, x_1 x_3, \ldots, x_{n-1} x_n]^\top$, and $\tilde{M}$ is a symmetric matrix of order $n + \frac{1}{2} n(n-1)$, then $\tilde{M}$ is not uniquely determined. The non-uniqueness follows from the identities:

$$(x_i x_j)^2 = (x_i^2)(x_j^2)$$
$$(x_i x_j)(x_i x_k) = (x_i^2)(x_j x_k)$$
$$(x_i x_j)(x_k x_l) = (x_i x_k)(x_j x_l) = (x_i x_l)(x_j x_k)$$

It is easy to see that the possible choices for $\tilde{M}$ define an affine space (see below for a closer description of that space).

2.1. SUM-OF-SQUARES DECOMPOSITIONS

Condition (5) will certainly hold if the polynomial P can be written as a sum of squares (s.o.s., in short), i.e., if

$$P(\mathbf{x}) = \sum_{i=1}^{t} f_i(\mathbf{x})^2$$

for some polynomial functions $f_i(\mathbf{x})$ $(i = 1, \ldots, t)$. A sum of squares decomposition is possible if and only if a representation of $P(\mathbf{x})$ exists where $\tilde{M}$ in (6) is positive semidefinite. We will show this for any homogeneous polynomial of degree $2r$ below, but introduce some convenient notation beforehand: for any $\mathbf{x} \in \mathbb{R}^n$ and any multi-index $\mathbf{m} \in \mathbb{N}_0^n$ (with $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$) we define $|\mathbf{m}| = \Sigma_i m_i$ and denote by $\mathbf{x}^{\mathbf{m}} = \Pi_i x_i^{m_i}$ the corresponding monomial of degree $|\mathbf{m}|$. Also, denote by $I^n(s) = \{\mathbf{m} \in \mathbb{N}_0^n : |\boldsymbol{m}| = s\}$ the set of all possible exponents of monomials of degree s (there are $d = \binom{n+s-1}{s}$ of them) and, as usual, $2I^n(s) = \{2\mathbf{m} : \mathbf{m} \in I^n(s)\}$. Finally, given a set of multi-indices I and a vector $\mathbf{x} \in \mathbb{R}^n$, we define $[\mathbf{x^m}]_{\mathbf{m}\in I}$ as the vector with components $\mathbf{x^m} = \Pi_i x_i^{m_i}$ for each $\mathbf{m} \in I$.

LEMMA 2.1. *If $\bar{P}(\mathbf{x})$ is a homogeneous polynomial of degree $2s$ in n variables $\mathbf{x} = [x_1, \ldots, x_n]^\top$, which has a representation*

$$\bar{P}(\mathbf{x}) = \sum_{i=1}^{l} f_i(\mathbf{x})^2$$

for some polynomials $f_i(\mathbf{x})$ $(i = 1, \ldots, l)$, then there are polynomials $h_i(\mathbf{x})$ which are homogeneous of degree s for all i such that $\bar{P}(\mathbf{x}) = \Sigma_{i=1}^{t} h_i(\mathbf{x})^2$ with $1 \leqslant t \leqslant l$. Further, $\bar{P}$ has a s.o.s. representation as above if and only if there is a symmetric positive-semidefinite matrix $d \times d$ matrix $\tilde{M} \in \mathcal{S}_d^+$ such that

$$\bar{P}(\mathbf{x}) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}} \tag{7}$$

where $d = \binom{n+s-1}{s}$ and $\tilde{\mathbf{x}} = [\mathbf{x^k}]_{\mathbf{k}\in I^n(s)} \in \mathbb{R}^d$.

Proof. It is easy to see that the degree of f_i is at most s for each i: if we assume to the contrary that f_j has maximal degree of all f_i and its degree exceeds s, then the square of its leading term which appears in the s.o.s. will never cancel out, since there can be no monomials of the same degree with negative coefficients.

We can therefore decompose each f_i as $f_i = h_i + g_i$ where h_i is homogeneous of degree s (or zero, but without loss of generality we assume that this happens only if $t < i \leqslant l$, including the possibility of $t = l$), and the degree of g_i is less than s. Now $\Sigma_{i=1}^{l} f_i(\mathbf{x})^2 = h(\mathbf{x}) + g(\mathbf{x})$ with $h(\mathbf{x}) = \Sigma_{i=1}^{t} h_i(\mathbf{x})^2$ and $g(\mathbf{x}) = \Sigma_{i=1}^{l} g_i(\mathbf{x})[2h_i(\mathbf{x}) + g_i(\mathbf{x})]$, so that g has degree less than $2s$ while h, as $\bar{P}$ itself, is homogeneous of degree $2s$. Thus $\bar{P} = h + g$ implies $g = 0$ and $\bar{P} = h$. Next note that any homogeneous h_i can be written as $h_i(\mathbf{x}) = \mathbf{a}_i^\top \tilde{\mathbf{x}}$ for some $\mathbf{a}_i \in \mathbb{R}^d$, so that $h(\mathbf{x}) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}}$ with $\tilde{M} = \Sigma_{i=1}^{t} \mathbf{a}_i \mathbf{a}_i^\top \in \mathcal{S}_d^+$. The converse is obvious via spectral decomposition of $\tilde{M}$. □

Next let us characterize all the matrices $\tilde{M} \in \mathcal{S}_d$ which allow for a representation $\bar{P}(\mathbf{x}) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}}$ for a given homogeneous polynomial $\bar{P}$.

LEMMA 2.2. *Let $\bar{P}(\mathbf{x}) = \Sigma_{\mathbf{m}\in I^n(s)} A_{\mathbf{m}} \mathbf{x}^{2\mathbf{m}}$ be a homogeneous polynomial of degree $2s$ in n variables $\mathbf{x} = [x_1, \ldots, x_n]^\top$ and define $\tilde{M} \in \mathcal{S}_d$ and $\tilde{\mathbf{x}} \in \mathbb{R}^d$ as in Lemma 2.1. Then $\bar{P}(\mathbf{x}) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}}$ if and only if*

$$\sum_{(\mathbf{j},\mathbf{k})\in[I^n(s)]^2:\mathbf{j}+\mathbf{k}=2\mathbf{m}} \tilde{M}_{\mathbf{j},\mathbf{k}} = A_{\mathbf{m}} \quad \textit{for all } \boldsymbol{m} \in I^n(s)\,, \tag{8}$$

$$\sum_{(\mathbf{j},\mathbf{k})\in[I^n(s)]^2:\mathbf{j}+\mathbf{k}=\mathbf{n}} \tilde{M}_{\mathbf{j},\mathbf{k}} = 0 \quad \textit{for all } \boldsymbol{n} \in I^n(2s)\backslash 2I^n(s)\,. \tag{9}$$

Proof. Observe that $\tilde{M}_{\mathbf{j},\mathbf{k}}\mathbf{x}^{\mathbf{j}}\mathbf{x}^{\mathbf{k}} = \tilde{M}_{\mathbf{j},\mathbf{k}}\mathbf{x}^{\mathbf{j}+\mathbf{k}}$. The assertion now follows by equating the corresponding coefficients of the polynomials $\tilde{\mathbf{x}}^\top \tilde{M}\tilde{\mathbf{x}}$ and $\sum_{\mathbf{m}\in I^n(s)} A_{\mathbf{m}}\mathbf{x}^{2\mathbf{m}}$. □

Parrilo showed [14] that $P(\mathbf{x})$ in (5) allows a sum of squares decomposition if and only if $M \in \mathcal{S}_n^+ + \mathcal{N}_n$, which is a well-known sufficient condition for copositivity. For completeness of presentation, we give a new proof here, which follows easily from the preceding lemmas.

Let us define the cone $\mathcal{K}_n^0 := \mathcal{S}_n^+ + \mathcal{N}_n = \mathcal{D}_n^*$, the cone dual dual to that of all doubly nonnegative matrices.

THEOREM 2.1 (Parrilo [14]). $P(\mathbf{x}) = (\mathbf{x}\circ\mathbf{x})^\top M(\mathbf{x}\circ\mathbf{x})$ *allows for a polynomial s.o.s. if and only if* $M \in \mathcal{K}_n^0$, *i.e., if and only if* $M = S + T$ *for matrices* $S \in \mathcal{S}_n^+$ *and* $T \in \mathcal{N}_n$.

Proof. Let $M \in \mathcal{S}_n$ and $P(\mathbf{x}) := (\mathbf{x}\circ\mathbf{x})^\top M(\mathbf{x}\circ\mathbf{x})$. In this case, the degree $2s$ of P equals four, so that $s = 2$, and hence $d = \binom{n+1}{2}$. Obviously, $A_{2\mathbf{e}_i} = M_{ii}$ while $A_{\mathbf{e}_i+\mathbf{e}_j} = 2M_{ij}$ if $1 \leqslant i < j \leqslant n$, by symmetry of M. Therefore Lemma 2.2 yields $\tilde{M}_{2\mathbf{e}_i,2\mathbf{e}_i} = M_{ii}$ while $\tilde{M}_{\mathbf{e}_i+\mathbf{e}_j,\mathbf{e}_i+\mathbf{e}_j} + \tilde{M}_{2\mathbf{e}_i,2\mathbf{e}_j} = 2M_{ij}$ if $1 \leqslant i < j \leqslant n$, since also $\tilde{M}$ is assumed to be symmetric. Note that we may and do assume that $\tilde{M}$ is positive-semidefinite by Lemma 2.1. Now put $T_{ij} = \frac{1}{2}\tilde{M}_{\mathbf{e}_i+\mathbf{e}_j,\mathbf{e}_i+\mathbf{e}_j}$ if $i \neq j$ while $T_{ii} = 0$, all i. Then $T \in \mathcal{N}_n$ because diagonal elements of $\tilde{M}$ cannot be negative. Further, $S = M - T$ satisfies $S_{ij} = \tilde{M}_{2\mathbf{e}_i,2\mathbf{e}_j}$ for all i, j, which means that S is a principal submatrix of the positive-semidefinite matrix $\tilde{M}$. Hence also $S \in \mathcal{S}_n^+$, which shows the necessity part of the assertion. To establish sufficiency, observe that $(\mathbf{x}\circ\mathbf{x})^\top S(\mathbf{x}\circ\mathbf{x})$ is, by spectral decomposition of S, even a s.o.s. in the variables $z_i = x_i^2$ while $(\mathbf{x}\circ\mathbf{x})^\top T(\mathbf{x}\circ\mathbf{x}) = \sum_{i,j}[\sqrt{T_{ij}}x_i x_j]^2$ is obviously a s.o.s. Hence $(\mathbf{x}\circ\mathbf{x})^\top M(\mathbf{x}\circ\mathbf{x})$ is, as the sum of two s.o.s. decompositions, itself a s.o.s. □

Higher order sufficient conditions can be derived by considering the polynomial:

$$P^{(r)}(\mathbf{x}) = P(\mathbf{x})\left(\sum_{k=1}^n x_k^2\right)^r = \sum_{i,j=1}^n M_{ij}x_i^2 x_j^2\left(\sum_{k=1}^n x_k^2\right)^r, \tag{10}$$

and asking when $P^{(r)}(\mathbf{x})$ has a sum of squares decomposition. It is clear from Lemma 2.1 that the set of matrices M which satisfy this condition forms a convex cone.

DEFINITION 2.1 (De Klerk and Pasechnik [5]). The convex cone $\mathcal{K}_n^r$ consists of the matrices for which $P^{(r)}(\mathbf{x})$ in (10) allows a polynomial sum of squares decomposition.

Obviously, these cones are contained in each other: $\mathcal{K}_n^r \subseteq \mathcal{K}_n^{r+1}$ for all r. This follows from

$$P^{(r+1)}(\mathbf{x}) = \sum_k x_k^2 P^{(r)}(\mathbf{x}) = \sum_{i,k} [f_i(\mathbf{x})x_k]^2 \,.$$

By explicitly calculating the coefficients $A_{\mathbf{m}}(M)$ of the homogeneous polynomial $P^{(r)}(\mathbf{x})$ of degree $2(r+2)$ and summarizing the above auxiliary results, we arrive at a characterization of $\mathscr{K}_n^r$ which has not appeared in the literature before.

THEOREM 2.2. *Let $n, r \in \mathbb{N}$ and $d = \binom{n+r+1}{r+2}$. Further, abbreviate $\mathbf{m}(i,j) = \mathbf{m} - \mathbf{e}_i - \mathbf{e}_j$ for any $\mathbf{m} \in \mathbb{R}^n$ and introduce the multinomial coefficients*

$$\begin{aligned} c(\mathbf{m}) &= |\mathbf{m}|! \Big/ \prod_i (m_i)! \,, && \text{if } \mathbf{m} \in \mathbb{N}_0^n \,, \\ c(\mathbf{m}) &= 0 \,, && \text{if } \mathbf{m} \in \mathbb{R}^n \setminus \mathbb{N}_0^n \,. \end{aligned} \tag{11}$$

For a symmetric matrix $M \in \mathscr{S}_n$, define

$$A_{\mathbf{m}}(M) = \sum_{i,j} c(\mathbf{m}(i,j)) M_{ij} \,. \tag{12}$$

Then $M \in \mathscr{K}_n^r$ if and only if there is a symmetric positive-semidefinite $d \times d$ matrix $\tilde{M} \in \mathscr{S}_d^+$ such that

$$\begin{aligned} \sum_{(\mathbf{j},\mathbf{k}) \in [I^n(r+2)]^2 : \mathbf{j}+\mathbf{k} = 2\mathbf{m}} \tilde{M}_{\mathbf{j},\mathbf{k}} &= A_{\mathbf{m}}(M) && \text{for all } \mathbf{m} \in I^n(r+2) \,, \\ \sum_{(\mathbf{j},\mathbf{k}) \in [I^n(r+2)]^2 : \mathbf{j}+\mathbf{k} = \mathbf{n}} \tilde{M}_{\mathbf{j},\mathbf{k}} &= 0 && \text{for all } \mathbf{n} \in I^n(2r+4) \setminus 2I^n(r+2) \,. \end{aligned} \tag{13}$$

Proof. By the multinomial law,

$$\begin{aligned} P^{(r)} &\equiv \sum_{i,j=1}^n M_{ij} x_i^2 x_j^2 \left(\sum_{k=1}^n x_k^2 \right)^r \\ &= \sum_{i,j=1}^n M_{ij} x_i^2 x_j^2 \sum_{\mathbf{k} \in I^n(r)} c(\mathbf{k}) \mathbf{x}^{2\mathbf{k}} \\ &= \sum_{\mathbf{k} \in I^n(r)} \sum_{i,j}^n c(\mathbf{k}) M_{ij} \mathbf{x}^{2\mathbf{k} + 2\mathbf{e}_i + 2\mathbf{e}_j} \\ &= \sum_{\mathbf{m} \in I^n(r+2)} \left[\sum_{i,j=1}^n c(\mathbf{m}(i,j)) M_{ij} \right] \mathbf{x}^{2\mathbf{m}} \,. \end{aligned} \tag{14}$$

The last identity follows by setting $\mathbf{m} = \mathbf{k} + \mathbf{e}_i + \mathbf{e}_j$. Hence $A_{\mathbf{m}}(M)$ as given by (12) are the coefficients of $P^{(r)}$, and the assertions follow by observing $s = r + 2$ with the help of Lemma 2.1 and Lemma 2.2. □

The following auxiliary result simplifies the expressions $A_{\mathbf{m}}(M)$ considerably.

LEMMA 2.3. *Let M be an arbitrary $n \times n$ matrix and denote by* $\operatorname{diag} M = [M_{ii}]_i \in$

$\mathbb{R}^n$ the vector obtained by extracting the diagonal elements of M. If $A_{\mathbf{m}}(M)$ is defined as in (12), then

$$A_{\mathbf{m}}(M) = \frac{c(\mathbf{m})}{s(s-1)} [\mathbf{m}^\top M\mathbf{m} - \mathbf{m}^\top \operatorname{diag} M] \quad \text{for all } \mathbf{m} \in I^n(s), \quad s \in \mathbb{N}. \quad (15)$$

Proof. Note that by definition, $c(\mathbf{m}(i,j)) = 0$ if $m_i m_j = 0$ in case $i \neq j$ while even $c(\mathbf{m}(i,i)) = 0$ if $m_i < 2$ so that nonzero coefficients of M_{ij} occur only for some (i,j) pairs depending on $\mathbf{m}$. Hence straightforward calculation shows, using $s = |\mathbf{m}|$,

$$\begin{aligned} A_{\mathbf{m}}(M) &= \sum_{i,j} c(\mathbf{m}(i,j)) M_{ij} \\ &= \sum_i \frac{c(\mathbf{m}) m_i (m_i - 1)}{s(s-1)} M_{ii} + \sum_{i \neq j} \frac{c(\mathbf{m}) m_i m_j}{s(s-1)} M_{ij} \\ &= \frac{c(\mathbf{m})}{s(s-1)} \left[\sum_i m_i^2 M_{ii} + \sum_{i \neq j} m_i m_j M_{ij} - \sum_i m_i M_{ii} \right] \end{aligned}$$

which exactly corresponds to (15). □

Observe that for $M = E_n$, we have, from $\mathbf{m}^\top E_n \mathbf{m} = (\mathbf{e}^\top \mathbf{m})^2 = |\mathbf{m}|^2$, thus

$$A_{\mathbf{m}}(E_n) = \frac{c(\mathbf{m})}{s(s-1)} [s^2 - s] = c(\mathbf{m}) \quad \text{for all } \mathbf{m} \in I^n(s), s \in \mathbb{N}. \quad (16)$$

Parrilo [14] showed that $M \in \mathcal{K}_n^1$ if the following system of linear matrix inequalities has a solution.

$$M - M^{(i)} \in \mathcal{S}_n^+, \quad i = 1, \ldots, n, \quad (17)$$

$$M_{ii}^{(i)} = 0, \quad i = 1, \ldots, n, \quad (18)$$

$$M_{ii}^{(j)} + 2M_{ij}^{(i)} = 0, \quad i \neq j, \quad (19)$$

$$M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} \geq 0, \quad i < j < k, \quad (20)$$

where $M^{(i)} \in \mathcal{S}_n$ for $i = 1, \ldots, n$.

The converse is also true: if $M \in \mathcal{K}_n^1$ then the system of LMI's (17)–(20) has a solution; this was used by De Klerk and Pasechnik [5] without giving a rigorous proof. We will now give a complete proof, by using our new characterizations of the cones $\mathcal{K}_n^r$ in Theorem 2.2 for $r = 1$. Note that $d = \binom{n+2}{3}$ in this case. We will use a shorthand notation where ijk as a subscript indicates the multi-index $\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k \in I^n(3)$.

THEOREM 2.3. *$M \in \mathcal{K}_n^1$ if and only if there are n symmetric $n \times n$ matrices $M^{(i)} \in \mathcal{S}_n$ for $i = 1, \ldots, n$ such that the system of LMI's (17)–(20) is satisfied.*

Proof. First assume that $M \in \mathscr{K}_n^1$. By Theorem 2.2 there exists a $\tilde{M} \in \mathscr{S}_d^+$ satisfying (13) such that

$$P^{(1)}(\mathbf{x}) \equiv \sum_{i,j}^{n} M_{ij} x_i^2 x_j^2 \left(\sum_{k=1}^{n} x_k^2 \right) = \tilde{\mathbf{x}}^\top \tilde{M} \tilde{\mathbf{x}} ,$$

where $\tilde{\mathbf{x}} + [\mathbf{x}^{\mathbf{k}}]_{\mathbf{k} \in I^n(3)} \in \mathbb{R}^d$, and $d = \binom{n+2}{3}$.

By (15), we have $A_{iii}(M) = M_{ii}$ while $A_{iij}(M) = 2M_{ij} + M_{ii}$ and $A_{ijk}(M) = 2[M_{ij} + M_{ik} + M_{jk}]$ if $1 \leq i < k \leq n$. Similarly, the left-hand sides of (13) read in case $\mathbf{n} = 2\mathbf{m}$

$$\begin{aligned}
&\tilde{M}_{iii,iii} , \quad \text{if } \mathbf{n} = 6\mathbf{e}_i , \\
&\tilde{M}_{iij,iij} + 2\tilde{M}_{iii,ijj} , \quad \text{if } \mathbf{n} = 4\mathbf{e}_i + 2\mathbf{e}_j , \quad i < j , \\
&\tilde{M}_{ijk,ijk} + 2[\tilde{M}_{iij,jkk} + \tilde{M}_{iik,jjk} + \tilde{M}_{ijj,ikk}] , \quad \text{if } \mathbf{n} = 2(\mathbf{e}_i + \mathbf{e}_j + \mathbf{e}_k) , \quad i < j < k .
\end{aligned} \tag{21}$$

Now put $S_{jk}^{(i)} = \tilde{M}_{ijj,ikk}$ for all triples (ijk). Then $S^{(i)} \in \mathscr{S}_n^+$ since it is a principal submatrix of the positive-semidefinite matrix $\tilde{M}$. Hence setting $M^{(i)} = M - S^{(i)}$ we see that condition (17) is satisfied. It remains to show that (18)–(20) hold. Now

$$M_{ii}^{(i)} = M_{ii} - S_{ii}^{(i)} = A_{iii}(M) - \tilde{M}_{iii,iii} = 0$$

and similarly

$$\begin{aligned}
M_{ii}^{(j)} + 2M_{ij}^{(i)} &= M_{ii} + 2M_{ij} - S_{ii}^{(j)} - 2S_{ij}^{(i)} \\
&= A_{iij}(M) - \tilde{M}_{iij,iij} - 2\tilde{M}_{iii,ijj} = 0 ,
\end{aligned}$$

whereas

$$\begin{aligned}
M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} &= M_{ij} + M_{ik} + M_{jk} - S_{jk}^{(i)} - S_{ik}^{(j)} - S_{ij}^{(k)} \\
&= \frac{1}{2} A_{ijk}(M) - \tilde{M}_{ijj,ikk} - \tilde{M}_{iij,jkk} - \tilde{M}_{iik,jjk} \\
&= \frac{1}{2} \tilde{M}_{ijk,ijk} \geq 0 ,
\end{aligned}$$

because the diagonal entries of $\tilde{M}$ cannot be negative. Thus we have constructed a solution to the system of LMI's (17)–(20).

Conversely, assume that a solution to (17)–(20) is given. Observe that

$$\begin{aligned}
P^{(1)}(\mathbf{x}) &= \sum_{i=1}^{n} x_i^2 (\mathbf{x} \circ \mathbf{x})^\top M (\mathbf{x} \circ \mathbf{x}) \\
&= \sum_{i=1}^{n} x_i^2 (\mathbf{x} \circ \mathbf{x})^\top (M - M^{(i)}(\mathbf{x} \circ \mathbf{x}) + \sum_{i=1}^{n} x_i^2 (\mathbf{x} \circ \mathbf{x})^\top (M^{(i)})(\mathbf{x} \circ \mathbf{x}) .
\end{aligned} \tag{22}$$

The first sum is obviously a s.o.s., since $M - M^{(i)} \in \mathscr{S}_n^+$ for every i. The second sum can likewise be written as a s.o.s. because of

$$\begin{aligned}\sum_i x_i^2 (\mathbf{x}\circ\mathbf{x})^\top M^{(i)}(\mathbf{x}\circ\mathbf{x}) &= \sum_{i,j,k} M_{jk}^{(i)} x_i^2 x_j^2 x_k^2 \\ &= \sum_i M_{ii}^{(i)}\, x_i^6 + \sum_{i\neq j} (M_{ii}^{(j)} + 2M_{ij}^{(i)}) x_i^4 x_j^2 \\ &\quad + \sum_{i<j<k} (2M_{jk}^{(i)} + 2M_{ik}^{(j)} + 2M_{ij}^{(k)}) x_i^2 x_j^2 x_k^2 \\ &= \sum_i (\sqrt{M_{ii}^{(i)}} x_i^3)^2 + \sum_{i\neq j} (\sqrt{M_{ii}^{(j)} + 2M_{ij}^{(i)}} x_i^2 x_j)^2 \\ &\quad + \sum_{i<j<k} (\sqrt{2[M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)}]} x_i x_j x_k)^2 , \end{aligned} \tag{23}$$

where we have used the non-negativity condition (20) to obtain the last equality. Note that the first two sums of the last expression vanish due to (18) and (19). Thus $P^{(1)}(\mathbf{x})$ is represented as a s.o.s. □

By closer inspection of the preceding proof we see that the condition (17) can be relaxed, to arrive at a (seemingly) less restrictive system of LMI's, namely:

$$M - M^{(i)} \in \mathcal{S}_n^+ + \mathcal{N}_n , \quad i = 1, \ldots, n , \tag{24}$$

$$M_{ii}^{(i)} = 0 , \quad i = 1, \ldots, n , \tag{25}$$

$$M_{jj}^{(i)} + 2M_{ij}^{(j)} = 0 , \quad i \neq j , \tag{26}$$

$$M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} \geqslant 0 , \quad i < j < k . \tag{27}$$

Indeed, the first sum in (22) is still a s.o.s., since $M - M^{(i)} \in \mathcal{S}_n^+ + \mathcal{N}_n = \mathcal{K}_n^0$ for every i, and because of Theorem 2.1. Hence (24)–(27) constitute an alternative characterization of $\mathcal{K}_n^1$, which in the next section will turn out to be quite insightful. There we will also specify an (apparently) even more relaxed characterization of $\mathcal{K}_n^1$, see (40)–(43) in Subsection 2.2 below. With slightly more effort, one could derive similar systems of LMIs for the cones $\mathcal{K}_n^r$ if $r \geqslant 2$. However, d then increases so rapidly with n (recall that $d = \mathcal{O}(n^{r+2})$) that the resulting problems become too large for current SDP solvers—even for small values of n.

We therefore change our perspective in the next subsection, to arrive at a series of LP approximations of the copositive cone. These approximations are weaker than the SDP ones, but can be solved more easily.

2.2. LP RELAXATIONS YIELDED BY NONNEGATIVITY

We start with a simple observation: If the polynomial $P^{(r)}(\mathbf{x})$ has only nonnegative coefficients, then it is already allows a sum-of-squares decomposition. This motivates the following definition.

DEFINITION 2.2 (De Klerk and Pasechnik [5]). The convex cone $\mathcal{C}_n^r$ consists of the matrices for which $P^{(r)}(\mathbf{x})$ in (10) has no negative coefficients. Hence for any r, we have $\mathscr{C}_n^r \subseteq \mathscr{K}_n^r$.

Again, we obviously have $\mathscr{C}_n^r \subseteq \mathscr{C}_n^{r+1}$ for all r. We can immediately derive a polyhedral representation of the cones $\mathscr{C}_n^r$; this characterization has not appeared in the literature.

THEOREM 2.4. *For any* $\mathbf{m} \in \mathbb{R}^n$, *define* $\operatorname{Diag} \mathbf{m}$ *as the* $n \times n$ *diagonal matrix containing* $\mathbf{m}$ *as its diagonal, i.e., satisfying* $\operatorname{diag}(\operatorname{Diag} \mathbf{m}) = \mathbf{m}$. *Then for all* $r \in \mathbb{N}_0$ *and* $n \in \mathbb{N}$,

$$\begin{aligned}\mathscr{C}_n^r &= \{M \in \mathscr{S}_n : \mathbf{m}^\top M \mathbf{m} - \mathbf{m}^\top \operatorname{diag} M \geqslant 0 \text{ for all } \mathbf{m} \in I^n(r+2)\} \\ &= \{M \in \mathscr{S}_n : \langle \mathbf{m}\mathbf{m}^\top - \operatorname{Diag} \mathbf{m}, M \rangle \geqslant 0 \text{ for all } \mathbf{m} \in I^n(r+2)\}\,.\end{aligned}$$

Proof. Follows from (14) in Theorem 2.2 and Lemma 2.3, with the help of the basic relations $\mathbf{m}^\top M \mathbf{m} = \langle \mathbf{m}\mathbf{m}^\top, M \rangle$ and $\mathbf{m}^\top \operatorname{diag} M = \langle \operatorname{Diag} \mathbf{m}, M \rangle$. □

Note that $\mathscr{C}_n^0 = \mathscr{N}_n$ since $I^n(2) = \{\mathbf{e}_i + \mathbf{e}_j : i, j\}$ while $M \in \mathscr{C}_n^1$ if and only if $M \in \mathscr{S}_n$ with

$$M_{ii} \geqslant 0\,, \quad i = 1, \ldots, n\,, \tag{28}$$

$$M_{ii} + 2M_{ij} \geqslant 0\,, \quad i \neq j\,, \tag{29}$$

$$M_{jk} + M_{ik} + M_{ij} \geqslant 0\,, \quad i < j < k\,. \tag{30}$$

This follows from Theorem 2.4 by the same arguments as in Theorem 2.3. We can also establish an alternative characterization of $\mathscr{C}_n^1$ similar to that in Theorem 2.3:

THEOREM 2.5. $M \in \mathscr{C}_n^1$ *and only if and only if there are* n *symmetric* $n \times n$ *matrices* $M^{(i)} \in \mathscr{S}_n$ *for* $i = 1, \ldots, n$ *such that the following system of linear inequalities has a solution*:

$$M - M^{(i)} \in \mathscr{N}_n\,, \quad i = 1, \ldots, n\,, \tag{31}$$

$$M_{ii}^{(i)} = 0\,, \quad i = 1, \ldots, n\,, \tag{32}$$

$$M_{ii}^{(j)} + 2M_{ij}^{(i)} = 0\,, \quad i \neq j\,, \tag{33}$$

$$M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} \geqslant 0\,, \quad i < j < k\,. \tag{34}$$

Proof. Suppose that $M \in \mathscr{C}_n^1$ and define $N_{jk}^{(i)}$ as follows:

$$\begin{aligned}&N_{jk}^{(i)} = M_{ii}\,, \quad \text{if } i = j = k\,, \\ &N_{jk}^{(i)} = \frac{1}{2} M_{ii} + M_{ij}\,, \quad \text{if } i = j \neq k \text{ or } i = k \neq j\,, \\ &N_{jk}^{(i)} = 0\,, \quad \text{else}\,.\end{aligned} \tag{35}$$

Then $N^{(i)} \in \mathcal{S}_n$ and because of (28) and (29) we get $N^{(i)} \in \mathcal{N}_n$. Further, $M^{(i)} = M - N^{(i)} \in \mathcal{S}_n$ satisfy $M_{ii}^{(i)} = 0$ and also

$$M_{ii}^{(j)} + 2M_{ij}^{(i)} = M_{ii} - 0 + 2\left[M_{ij} - \frac{1}{2} M_{ii} - M_{ij} \right] = 0 .$$

Finally, (30) implies (34) because $M_{jk}^{(i)} = M_{jk}$ if $\{i, j, k\}$ contains three distinct elements due to the definition of $N^{(i)}$. The converse follows as in the proof of Theorem 2.3, without taking square roots in (23).

By comparing (31)–(34) to (24)–(27) we see that merely $\mathcal{S}_n^+ + \mathcal{N}_n$ in (24) has been shrunk to $\mathcal{N}_n$ in (31). This reflects the fact that $\mathcal{C}_n^1 \subset \mathcal{K}_n^1$.

Further, the two equalities (32) and (33) can be replaced with inequalities, without changing the characterization: $M \in \mathcal{C}_n^1$ if and only if there are n symmetric $n \times n$ matrices $M^{(i)} \in \mathcal{S}_n$ for $i = 1, \ldots, n$ such that the following system of linear inequalities has a solution:

$$M - M^{(i)} \in \mathcal{N}_n , \quad i = 1, \ldots, n , \tag{36}$$

$$M_{ii}^{(i)} \geqslant 0 , \quad i = 1, \ldots, n , \tag{37}$$

$$M_{ii}^{(j)} + 2M_{ij}^{(i)} \geqslant 0 , \quad i \neq j , \tag{38}$$

$$M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} \geqslant 0 , \quad i < j < k . \tag{39}$$

Similarly, also $M \in \mathcal{K}_n^1$ if and only if there are n symmetric $n \times n$ matrices $M^{(i)} \in \mathcal{S}_n$ for $i = 1, \ldots, n$ such that the following system of LMIs has a solution:

$$M - M^{(i)} \in \mathcal{S}_n^+ + \mathcal{N}_n , \quad i = 1, \ldots, n , \tag{40}$$

$$M_{ii}^{(i)} \geqslant 0 , \quad i = 1, \ldots, n , \tag{41}$$

$$M_{jj}^{(i)} + 2M_{ij}^{(j)} \geqslant 0 , \quad i \neq j , \tag{42}$$

$$M_{jk}^{(i)} + M_{ik}^{(j)} + M_{ij}^{(k)} \geqslant 0 , \quad i < j < k . \tag{43}$$

Indeed, we may use non-negativity via (41), (42) and (43) in (23) to obtain the desired s.o.s. decomposition there, and an analogous argument without taking square roots applies to establish sufficiency of (36)–(39).

Every strictly copositive matrix M lies in some cone $\mathcal{C}_n^r$ for r sufficiently large; this follows from a famous theorem of Pólya [16] (see also Powers and Reznick [17]). In summary, we have the following theorem.

THEOREM 2.6 (De Klerk and Pasechnik [5]). *Let $M \notin \mathcal{S}_n^+ + \mathcal{N}_n$ be strictly copositive. Then there are integers $r_{\mathcal{K}}(M)$ and $r_{\mathcal{C}}(M)$ with $1 \leqslant r_{\mathcal{K}}(M) \leqslant r_{\mathcal{C}}(M) < +\infty$, such that*

$$\mathcal{S}_n^+ + \mathcal{N}_n = \mathcal{K}_n^0 \subset \mathcal{K}_n^1 \subset \cdots \subset \mathcal{K}_n^r \ni M$$

for all $r \geq r_{\mathscr{K}}(M)$ *while* $M \notin \mathscr{K}^{r_{\mathscr{K}}(M)-1}$, *and similarly*

$$\mathscr{N}_n = \mathscr{C}_n^0 \subset \mathscr{C}_n^1 \subset \cdots \mathscr{C}_n^r \ni M$$

for all $r \geq r_{\mathscr{C}}(M)$ *while* $M \notin \mathscr{C}_n^{r_{\mathscr{C}}(M)-1}$.

The first part of the theorem (concerning the cones $\mathscr{K}_n^r$) already follows from arguments by Parrilo [14].

3. Approximation results

In this section we consider families of LP and SDP approximations to $p^* = \min\{\mathbf{z}^\top Q\mathbf{z} : \mathbf{z} \in \Delta\}$, and prove a bound on the quality of these approximations.

3.1. LP-BASED APPROXIMATIONS

Let us define:

$$\mathrm{p}_{\mathscr{C}}^{(\mathrm{r})} = \min\{\langle Q, X\rangle : \langle E_n, X\rangle = 1, X \in (\mathscr{C}_n^r)^*\}, \tag{44}$$

for $r = 0, 1, \ldots$ which has dual formulation

$$p_{\mathscr{C}}^{(r)} = \max\{\lambda : Q - \lambda E_n \in \mathscr{C}_n^r, \lambda \in \mathbb{R}\}. \tag{45}$$

Note that problem (45) is a relaxation of problem (4) where the copositive cone is approximated by $\mathscr{C}_n^r$. It therefore follows that $p_{\mathscr{C}}^{(r)} \leq p^*$ for all r. We now provide an alternative representation of $p_{\mathscr{C}}^{(r)}$. This representation uses the following rational grid which approximates the standard simplex:

$$\Delta(r) = \frac{1}{r+2} I^n(r+2) = \{\mathbf{y} \in \Delta : (r+2)\mathbf{y} \in \mathbb{N}_0^n\}. \tag{46}$$

A naive approximation of problem (1) would be

$$p_{\Delta(r)} := \min\{\mathbf{y}^\top Q\mathbf{y} : \mathbf{y} \in \Delta(r)\} \geq p^*. \tag{47}$$

The next theorem shows that there is a close link between $p_{\mathscr{C}}^{(r)}$ and the naive approximation $p_{\Delta(r)}$. In particular, one can obtain $p_{\mathscr{C}}^{(r)}$ in a similar way as the naive approximation $p_{\Delta(r)}$ is obtained, i.e., by only doing function evaluations at points on the grid $\Delta(r)$.

THEOREM 3.1. *For any* $r \in \mathbb{N}_0$ *consider the rational discretization* $\Delta(r)$ *of the standard simplex* Δ *from (46). If* Q *is an arbitrary symmetric* $n \times n$ *matrix and* $\mathbf{q}_r = \frac{1}{r+2} \operatorname{diag} Q$, *then*

$$p_{\mathscr{C}}^{(r)} = \frac{r+2}{r+1} \min\{\mathbf{y}^\top Q\mathbf{y} - \mathbf{q}_r^\top \mathbf{y} : \mathbf{y} \in \Delta(r)\}. \tag{48}$$

Proof. First we use the representation of $\mathscr{C}_n^r$ from Theorem 2.4, putting $M = Q - \lambda E_n$, and observing that from (15) and (16),

$$A_{\mathbf{m}}(M) = A_{\mathbf{m}}(Q) - \lambda A_{\mathbf{m}}(E_n) = c(\mathbf{m})\left[\frac{1}{(r+2)(r+1)}(\mathbf{m}^\top Q\mathbf{m} - \mathbf{m}^\top \operatorname{diag} Q) - \lambda\right].$$

Then by (45) we have

$$p_{\mathscr{C}}^{(r)} = \min\left\{\frac{1}{(r+2)(r+1)}(\mathbf{m}^\top Q\mathbf{m} - \mathbf{m}^\top \operatorname{diag} Q) : \mathbf{m} \in I^n(r+2)\right\},$$

which gives (48) by putting $\mathbf{y} = \frac{1}{r+2}\mathbf{m} \in \Delta(r)$. □

Observe that, compared to the naive approximation $p_{\Delta(r)}$, we subtract a linear correction term $\mathbf{q}_r^\top \mathbf{y}$ from the original objective $\mathbf{y}^\top Q\mathbf{y}$ to get closer to p^*, but we have to compensate with a factor $(r+2)/(r+1) = 1 + \mathscr{O}(1/r) > 1$, because $p_{\mathscr{C}}^{(r)}$ always exceeds p^*. Given the last result, it is straightforward to derive the following approximation guarantee.

THEOREM 3.2. *Let* $\bar{p} := \max_{\mathbf{x}\in\Delta} \mathbf{x}^\top Q\mathbf{x}$. *One has*

$$p^* - p_{\mathscr{C}}^{(r)} \leqslant \frac{1}{r+1}(\bar{p} - p^*)$$

as well as

$$p_{\Delta(r)} - p^* \leqslant \frac{1}{r+2}(\bar{p} - p^*).$$

Proof. By Theorem 3.1 we have

$$\begin{aligned} p_{\mathscr{C}}^{(r)} &= \frac{r+2}{r+1}\min\{\mathbf{y}^\top Q\mathbf{y} - \mathbf{q}_r^\top \mathbf{y} : \mathbf{y} \in \Delta(r)\} \\ &\geqslant \frac{r+2}{r+1}\left(p^* - \max_{\mathbf{y}\in\Delta(r)} \frac{1}{r+2}(\operatorname{diag} Q)^T\mathbf{y}\right) \\ &= \frac{r+2}{r+1}\left(p^* - \frac{1}{r+2}\max_i Q_{ii}\right) \\ &\geqslant \frac{r+2}{r+1}\left(p^* - \frac{1}{r+2}\bar{p}\right) \\ &= p^* + \frac{1}{r+1}(p^* - \bar{p}). \end{aligned}$$

The first result follows. The second relation is derived in a similar way: by Theorem 3.1 we have

$$\begin{aligned} \min_{\mathbf{y}\in\Delta(r)} \mathbf{y}^\top Q\mathbf{y} &\leqslant \frac{r+1}{r+2}p_{\mathscr{C}}^{(r)} + \frac{1}{r+2}\max_i Q_{ii} \\ &\leqslant \frac{r+1}{r+2}p^* + \frac{1}{r+2}\bar{p}, \end{aligned}$$

which implies the second statement. □

3.2. SDP-BASED APPROXIMATIONS

Similarly as in the definition of $p^{(r)}_{\mathcal{C}}$, we can define SDP-based approximations to p^* using the cones $\mathcal{K}^r_n$ instead of $\mathcal{C}^r_n$, namely:

$$p^{(r)}_{\mathcal{K}} = \min\{\langle Q, X\rangle : \langle E_n, X\rangle = 1, X \in (\mathcal{K}^r_n)^*\}\,, \tag{49}$$

for $r = 0, 1, \ldots$. The dual problem of (49) is

$$\max\{\lambda : Q - \lambda E_n \in \mathcal{K}^r_n,\ \lambda \in \mathbb{R}\}\,. \tag{50}$$

It may not immediately be clear that the optimal values in (49) and (50) are attained and equal. Hence we prove the following result.

THEOREM 3.3. *The problem pair (49) and (50) have the same optimal value, namely $p^{(r)}_{\mathcal{K}}$, and both problems attain this optimal value. In particular,*

$$p^{(r)}_{\mathcal{K}} = \max\{\lambda : Q - \lambda E_n \in \mathcal{K}^r_n,\ \lambda \in \mathbb{R}\}\,. \tag{51}$$

Proof. In order to invoke Theorem 1.1, we have to show that there is a matrix X in the relative interior of the feasible region of problem (49), and that there is a feasible solution $\lambda \in \mathbb{R}$ of (50) such that $Q - \lambda E_n$ is in the interior of $\mathcal{K}^r_n$. To establish the latter property, simply take $\lambda = -1$ which defines a matrix $Q + E_n$ in the interior of $\mathcal{N}_n + \mathcal{S}^+_n = \mathcal{K}^0_n$. Indeed, we assumed right from the start that Q has no negative entries, so the same holds true for any sufficiently small perturbation of $Q + E_n$. Hence this perturbation lies also in $\mathcal{N}_n \subset \mathcal{K}^0_n \subset \mathcal{K}^r_n$. Consequently, $Q + E_n$ lies in the interior of $\mathcal{K}^r_n$ for all $r = 0, 1, \ldots$. We proceed to establish strict feasibility of the matrix $X = \frac{1}{n^2+n}(nI_n + E_n)$ which clearly satisfies $\langle E_n, X\rangle = 1$. Consider the symmetric square-root factorization $W = \sqrt{X}$ of X, which is given by $W = 1/(\sqrt{n+1})I_n + (\sqrt{2}-1)/(n\sqrt{n+1})E_n$ (it is straightforward to verify that $W^2 = X$). Obviously, all entries of W are strictly positive, hence the same holds true for the symmetric square-root factorization $\sqrt{U}$ of any sufficiently small perturbation U of X, by continuity of the map $U \mapsto \sqrt{U}$ at the positive-definite matrix X. But since $\sqrt{U}$ has no negative entries, we conclude that $U \in \mathcal{C}^*_n \subset (\mathcal{K}^r_n)^*$. Therefore X lies in the interior of $(\mathcal{K}^r_n)^*$ for all $r = 0, 1, \ldots$, and Theorem 1.1 establishes the desired strong duality assertion. □

COROLLARY 3.1. *For $p^{(r)}_{\mathcal{K}}$ as in (51), we have $p^{(r)}_{\mathcal{K}} \geqslant p^{(r)}_{\mathcal{C}}$ for all $r = 0, 1, \ldots$, and therefore*

$$p^* - p^{(r)}_{\mathcal{K}} \leqslant \frac{1}{r+1}(\bar{p} - p^*)\,.$$

Proof. The first inequality follows from $\mathcal{C}^r_n \subset \mathcal{K}^r_n$, and the second from the first and from Theorem 3.2. □

4. Comparison with known approximation results

4.1. AN IMPLEMENTABLE POLYNOMIAL-TIME APPROXIMATION SCHEME

Consider the generic optimization problem:

$$\phi^* := \max\{f(\mathbf{x}) : \mathbf{x} \in S\},$$

for some nonempty, bounded convex set S, and let

$$\phi_* := \min\{f(\mathbf{x}) : \mathbf{x} \in S\}.$$

DEFINITION 4.1 (see, e.g., [12]). A value ψ is said to approximate ϕ_* with relative accuracy $\mu \in [0, 1]$ if

$$|\psi - \phi_*| \leq \mu(\phi^* - \phi_*).$$

The approximation is called implementable if $\psi \geq \phi_*$.

Note that ψ is an *implementable* approximation if and only if $\psi = f(\mathbf{x})$ for some $\mathbf{x} \in S$.

DEFINITION 4.2 (see, e.g., [13]). If a class of optimization problems allows an implementable, polynomial-time μ-approximation for each $\mu > 0$, then we say that this problem class allows a polynomial-time approximation scheme (PTAS).

It is known from Bellare and Rogaway [1] that (even in a weaker sense) there is no polynomial-time μ-approximation of the optimal value of the problem

$$\min\{\mathbf{x}^\top Q\mathbf{x} : B\mathbf{x} = \mathbf{b}, \mathbf{o} \leq \mathbf{x} \leq \mathbf{e}\} \tag{52}$$

for some $\mu \in (0, \frac{1}{3})$, unless $P = NP$.

Using semidefinite programming techniques, the problem

$$\min\left\{\mathbf{x}^\top Q\mathbf{x} : B\mathbf{x} = \frac{1}{2}\mathbf{e},\ \mathbf{o} \leq \mathbf{x} \leq \mathbf{e}\right\}$$

can be approximated in polynomial-time with the relative accuracy $(1 - \mathcal{O}(1/\log(n)))$, see [12] and the references therein.

Note the standard quadratic optimization problem (1) is a special case of this problem where $B = (1/2)\mathbf{e}^\top$. Nesterov [10] has shown that problem (1) allows a polynomial-time, implementable 2.3-approximation. Our result in Theorem 3.2 improves on this result, since it can be restated as: $p_{\Delta(r)}$ is a polynomial-time, implementable, $(1/(r+2))$-approximation of p^*. In other words, for any given $\epsilon > 0$ we obtain an implementable polynomial-time ϵ-approximation for problem (1), i.e., a PTAS.

An intuitive explanation why standard QPs admit a PTAS is that the (relative) volume of the standard simplex decreases exponentially fast with increasing n, contrasting with the general case (52) treated by Bellare and Rogaway.

4.2. SPECIAL CASE: THE MAXIMUM STABLE SET PROBLEM

Let $G = (V, E)$ be a simple graph with $|V| = n$, and let $\alpha(G)$ denote its stability number (cardinality of the largest stable set). It is known from Motzkin-Straus [8] (see also De Klerk and Pasechnik [5]) that

$$\frac{1}{\alpha(G)} = \min_{\mathbf{x} \in \Delta} \mathbf{x}^\top (A + I_n)\mathbf{x}\,, \tag{53}$$

where A is the adjacency matrix of G. Note that this is a special case of the standard quadratic optimization problem (1), where $Q = A + I_n$.

The stability number $\alpha(G)$ cannot be approximated in polynomial-time to within a factor $|V|^{1/2-\epsilon}$ for any $\epsilon > 0$ unless $P = NP$, or within a factor $|V|^{1-\epsilon}$ for any $\epsilon > 0$ unless $NP = ZPP$ [7]. The complexity result in Theorem 3.2 guarantees that we obtain a μ-approximation of $1/\alpha(G)$ in polynomial-time for each $\mu > 0$. This does not contradict the above-mentioned in-approximability results—a μ-approximation of $1/\alpha(G)$ for a fixed $\mu \in (0, 1)$ can yield an arbitrarily bad approximation of $\alpha(G)$ (if $\alpha(G)$ is much larger than $1/\mu$).

THEOREM 4.1. *Define $\alpha^{(r)} = 1/p_{\mathscr{C}}^{(r)}$ where $p_{\mathscr{C}}^{(r)}$ is defined in (44). Assume without loss of generality that $\alpha^{(r)} \geqslant 2$ (i.e., assume G is not a complete graph). Then*

$$\lfloor \alpha^{(r)} \rfloor = \alpha(G) \quad \text{if and only if} \quad r \geqslant \alpha(G)^2 - 1\,. \tag{54}$$

Proof. To abbreviate, put $s = \alpha(G)$. The assumption $\alpha^{(r)} \geqslant 2$ means that $p_{\mathscr{C}}^{(r)} \leqslant \frac{1}{2}$. Now we apply Theorem 3.2, to arrive at

$$\frac{1}{s} - \frac{1}{\alpha^{(r)}} \leqslant \frac{1 - \frac{1}{s}}{r+1}\,. \tag{55}$$

Isolating $\alpha^{(r)}$ in (55) we can rewrite this inequality in equivalent form:

$$\alpha^{(r)} - s \leqslant \frac{s^2 - s}{r + 2 - s}\,.$$

Hence we find that $\alpha^{(r)} - s < 1$ if $r > s^2 - 2$. In other words, $\lfloor \alpha^{(r)} \rfloor = s$ if $r \geqslant s^2 - 1$. To show the converse, consider an independent set S of cardinality s. First suppose that $r + 2 = st$ for some integer t, and put $\mathbf{y} = 1/(r+2)[\Sigma_{i \in S}\, t\mathbf{e}_i] \in \Delta(r)$. Then straightforward calculations show $\mathbf{y}^\top Q\mathbf{y} - \mathbf{q}_r^\top \mathbf{y} = (t-1)/(r+2)$. Thus Theorem 3.1 entails

$$\frac{t-1}{r+1} \geqslant p_{\mathscr{C}}^{(r)}\,. \tag{56}$$

Now, if $\lfloor \alpha^{(r)} \rfloor \neq s$, then $p_{\mathscr{C}}^{(r)} = 1/\alpha^{(r)} > 1/(s+1)$, so that (56) gives $(s+1)(t-1) >$

$r+1=st-1$, yielding $t \geqslant s+1$ or $r+2 \geqslant s^2+s$. If, on the other hand, s does not divide $r+2$, then there are unique integers t, p such that $r+2=st+p$ with $1 \leqslant p \leqslant s-1$. By the monotonicity of the approximation, $\lfloor \alpha^{(r)} \rfloor = s$ implies $\lfloor \alpha^{(s(t+1))} \rfloor = s$, which, as shown above, gives $t+1 \geqslant s+1$. This establishes $r+2 = st+p \geqslant s^2+p$ or $r \geqslant s^2-1$. □

A slightly weaker version of the sufficiency part in (54) has been proven by De Klerk and Pasechnik [5]: if $r \geqslant \alpha(G)^2$, then $\lfloor \alpha^{(r)} \rfloor = \alpha(G)$. Further, setting $t=1$ in (56) and exploiting monotonicity again, we conclude that for any $r \leqslant \alpha(G)-2$, the LP-approximation $\alpha^{(r)}$ of $\alpha(G)$ is necessarily trivial in the sense that $p^{(r)}_{\mathscr{C}} = 0$. Thus, $p^{(r)}_{\mathscr{C}} > 0$ implies that $r \geqslant \alpha(G)-1$. Also this last result was already shown by De Klerk and Pasechnik [5]. This inequality, read the other way round, can be seen as an instant upper bound for the stability number $\alpha(G)$.

5. Examples

Here we give some examples for problems of various origin.

EXAMPLE 5.1. Consider an instance of the standard quadratic optimization problem (1), where the matrix Q is given by:

$$Q = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

This example corresponds to computation of the largest stable set in a pentagon (see Section 4.2 and [5]). We have $p^{(0)}_{\mathscr{C}} = 0$, $p^{(1)}_{\mathscr{C}} = \frac{1}{3} > 0$, giving the instant upper bound $\alpha(G) \leqslant 2$ which in this case is already exact. Passing to the SDP relaxation, we get $p^{(0)}_{\mathscr{K}} = 1/\sqrt{5} \approx 0.44721$ and finally $p^{(1)}_{\mathscr{K}} = p^* = \frac{1}{2}$. The proof that $p^{(1)}_{\mathscr{K}} = \frac{1}{2}$ requires the observation that the matrix

$$Q - \frac{1}{2} E_n = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{bmatrix}$$

is known to be in $\mathscr{K}^1_n$ (but it is not in $\mathscr{K}^0_n = \mathscr{S}^+_n + N_n$); for a proof, see [14].

EXAMPLE 5.2. Consider an instance of the standard quadratic optimization problem (1), where the matrix Q is given by:

$$Q = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

This example corresponds to computation of the largest stable set in the complement of the graph of an icosahedron (see Section 4.2 and [5]). Here $p^* = \frac{1}{3}$ while $p_{\mathcal{K}}^{(1)} \approx 0.309$ but the LP approximation yields the trivial lower bound: $p_{\mathcal{C}}^{(1)} = 0$, since $1 \leqslant \alpha(G) - 2$. This example shows that—even though the approximation $p_{\mathcal{K}}^{(1)}$ is much more expensive to compute than $p_{\mathcal{C}}^{(1)}$ (see Section 6)—it can yield a much better lower bound on p^*. This example—first considered in [5]—is the smallest problem we know where the $p_{\mathcal{K}}^{(1)}$ approximation to p^* is not exact. It remains an open problem to find the smallest value of n where the $p_{\mathcal{K}}^{(1)}$ approximation to p^* is not exact.

EXAMPLE 5.3. This example is from a mathematical model in population genetics [21], where, among others, the following matrix 410A is considered:

$$Q = \begin{bmatrix} 14 & 15 & 16 & 0 & 0 \\ 15 & 14 & 12.5 & 22.5 & 15 \\ 16 & 12.5 & 10 & 26.5 & 16 \\ 0 & 22.5 & 26.5 & 0 & 0 \\ 0 & 15 & 16 & 0 & 14 \end{bmatrix}.$$

and the objective is to maximize $\mathbf{x}^\top \bar{Q} \mathbf{x}$ subject to $\mathbf{x} \in \Delta$. There are five different local solutions to this problem. The globally optimal value here is $16\frac{1}{3}$ which corresponds to $\mathbf{x} = [0, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 0]^\top$. After changing to a minimization problem of the form (1) with a nonnegative coefficient matrix, we obtain the upper bound 21 for the optimal value via computation of $p_{\mathcal{C}}^{(1)}$, while the approximation via computation of $p_{\mathcal{K}}^{(1)}$ is exact.

EXAMPLE 5.4. This example deals with portfolio optimization and is taken from Berkelaar et al. [2]. Here, $\mathbf{x} \in \Delta$ corresponds to a portfolio: x_i is the fraction of your capital to be invested in investment i. Given a portfolio $\mathbf{x} \in \Delta$ there is a risk $\mathbf{x}^\top \bar{Q} \mathbf{x}$ associated with the portfolio which should be minimized, and an expected return $\mathbf{r}^\top \mathbf{x}$ to be maximized. An example from [2] has data:

$$\bar{Q} = \begin{bmatrix} 0.82 & -0.23 & 0.155 & -0.013 & -0.314 \\ -0.23 & 0.484 & 0.346 & 0.197 & 0.592 \\ 0.155 & 0.346 & 0.298 & 0.143 & 0.419 \\ -0.013 & 0.197 & 0.143 & 0.172 & 0.362 \\ -0.314 & 0.592 & 0.419 & 0.362 & 0.916 \end{bmatrix} \quad \text{and} \quad \mathbf{r} = \begin{bmatrix} 1.78 \\ 0.37 \\ 0.237 \\ 0.315 \\ 0.49 \end{bmatrix}.$$

(note that the matrix $\bar{Q}$ is *not* positive semidefinite). We can formulate this (multi-objective) problem in the form (1) as follows:

$$\min_{\mathbf{x} \in \leqslant} [\mathbf{x}^\top \bar{Q}\mathbf{x} - c(\mathbf{r}^\top \mathbf{x})^2] \,,$$

for some parameter $c > 0$ measuring risk-aversion; this problem is now of the form (1) if we set $\hat{Q} = \bar{Q} - c\mathbf{r}\mathbf{r}^\top$. To avoid negative entries, we replace $\hat{Q}$ with $Q = \hat{Q} + \gamma E_n$, where $\gamma = 0.4012$ is the maximal entry of $-\hat{Q}$. For this Q and $c = 0.1$, we have $p_{\mathcal{K}}^{(1)} \approx 0.4839$ and $p_{\mathcal{C}}^{(1)} \approx 0.3015$. Since $\mathbf{x} = [0.37, 0.26, 0, 0.37, 0]^\top \in \Delta$ yields the objective value 0.4839, this suggests that the SDP relaxation is exact.

6. Numerical results

We first compare the LP approximation $p_{\mathcal{C}}^{(1)}$ of p^* with the stronger approximation $p_{\mathcal{K}}^{(1)}$.

To this end, we generated 1000 instances of (1) with $n = 10$, and where the matrices Q are symmetric with entries uniformly distributed in [0, 1].

In Figure 1 we show a histogram with the 'distribution' of the ratios $p_{\mathcal{C}}^{(1)} : p_{\mathcal{K}}^{(1)}$ for the 1000 test problems. Note that in about $\frac{3}{4}$ of the cases the ratio is (close to) unity. This shows that the LP-based approximations are comparable to the SDP ones in this sense. However, the computation of $p_{\mathcal{C}}^{(1)}$ is much cheaper than that of $p_{\mathcal{K}}^{(1)}$—in Table 1 some typical solution times are given for growing values of n. In all cases the computer used was a Pentium II (450 MHz). The LP and SDP solver was SeDuMi [20] by Jos Sturm running under Matlab 5.3.

The next set of experiments we performed was to approximate the stability number of random graphs by computing the $p_{\mathcal{K}}^{(1)}$ and $p_{\mathcal{C}}^{(1)}$ approximations to the optimal value of problem (53). For this purpose we generated 20 graphs on 12 vertices to have stability number 6. The edges outside the maximal stable set were generated at random with the probability of including an edge between two given vertices outside the maximal stable set being $\frac{1}{2}$. In all cases the $p_{\mathcal{K}}^{(1)}$ approaching was exact while the $p_{\mathcal{C}}^{(1)}$ approximation gave the trivial zero lower bound (see the last remark of Section 4.2).

This indicates that Example 5.2 is quite special—it is difficult to find a graph on 12 vertices where $p_{\mathcal{K}}^{(1)}$ does not give the stability number.

7. Conclusions

We have suggested a novel approach to approximating the optimal value of the

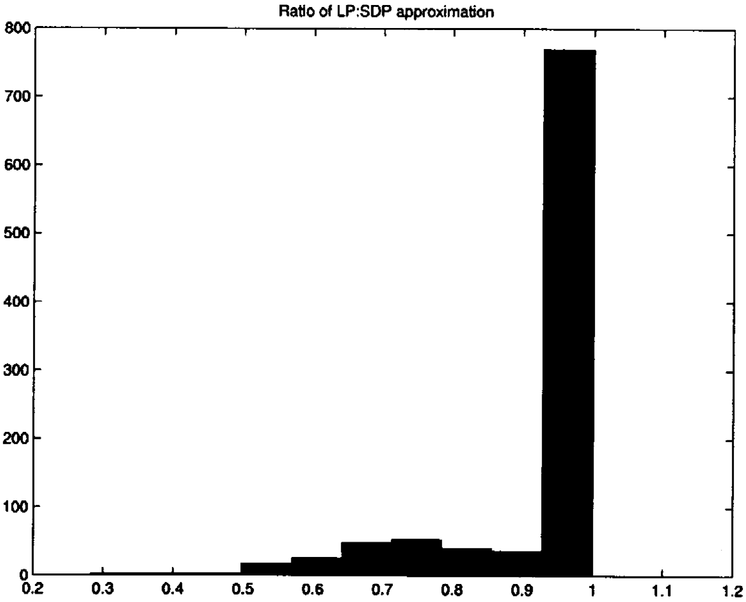

Figure 1. Histogram showing the distribution of the ratio $p_{\mathscr{C}}^{(1)} : p_{\mathscr{K}}^{(1)}$ approximation values for 1000 random instances of (1) with $n = 10$.

standard quadratic optimization problem (1)—the problem (1) is first rewritten as a copositive programming problem, and subsequently the copositive cone is approximated by using systems of linear or linear matrix inequalities. The resulting approximations are therefore LPs or SDPs, depending on which approximation scheme is used. Higher order approximations are also possible, where the resulting LPs and SDPs become larger, but the approximation is also provably better. In particular, we have quantified the quality of the approximation as a function of the order of approximation (Theorem 3.2). In particular, we have shown that our approximation is a polynomial-time approximation scheme (PTAS) of the standard quadratic optimization problem (1). Thus we have improved on the previously best known approximation result due to Nesterov [10].

Table 1. Typical CPU times for solving the respective LP and SDP relaxations for growing n

n	CPU time for $p_{\mathscr{C}}^{(1)}$ (s)	CPU time for $p_{\mathscr{K}}^{(1)}$ (s)
10	0.27	2.86
15	0.88	69.7
20	1.92	80.1
25	4.12	not run
30	6.86	not run
35	11.2	not run
40	26.5	not run

Moreover, we have presented numerical evidence showing the quality of the approximations.

Acknowledgement

The authors would like to thank Kees Roos for his comments on a draft version of this paper. The second author thanks Arkadi Nemirovski and Dima Pasechnik for useful discussions.

References

[1] Bellare, M. and Rogaway, P. (1995), The Complexity of Approximating a Nonlinear Program, *Mathematical Programming* 69, 429–441.
[2] Berkelaar, A.B., Jansen, B., Roos, C. and Terlaky, T. (1996), Sensitivity Analysis in (Degenerate) Quadratic Programming. Technical Report TWI 96-26, Reports of the Faculty of Technical Mathematics and Informatics, Delft University of Technology.
[3] Bomze, I.M. (1998), On Standard Quadratic Optimization Problems, *Journal of Global Optimization* 13, 369–387.
[4] Bomze, I.M., Dür, M., de Klerk, E., Roos, C., Quist, A. and Terlaky, T. (2000), On Copositive Programming and Standard Quadratic Optimization Problems, *Journal of Global Optimization* 18, 301–320.
[5] de Klerk, E. and Pasechnik, D.V. (2002), Approximation of the Stability Number of a Graph via Copositive Programming, *SIAM Journal of Optimization* 12(4), 875–892.
[6] Goemans, M.X. and Williamson, D.P. (1995), Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming, *Journal of the ACM* 42, 1115–1145.
[7] Hastad, J. (1999), Clique is Hard to Approximate Within $|V^{1-\epsilon}|$, *Acta Mathematica* 182, 105–142.
[8] Motzkin, T.S. and Straus, E.G. (1965), Maxima for Graphs and a New Proof of a Theorem of Túran, *Canadian J. Math.* 17, 533–540.
[9] Murty, K.G. and Kabadi, S.N. (1987), Some NP-Complete Problems in Quadratic and Linear Programming, *Mathematical Programming* 39, 117–129.
[10] Nesterov, Y.E. (1999), Global Quadratic Optimization on the Sets with Simplex Structure. Discussion paper 9915, CORE, Katholic University of Louvain, Belgium.
[11] Nesterov, Y.E. and Nemirovskii, A.S. (1994), *Interior Point Methods in Convex Programming: Theory and Applications*. SIAM, Philadelphia, PA.
[12] Nesterov, Y.E., Wolkowicz, H. and Ye, Y. (2000), Nonconvex Quadratic Optimization, in Wolkowicz, H., Saigal, R. and Vandenberghe, L. (eds.), *Handbook of Semidefinite Programming*, pp. 361–416. Kluwer Academic Publishers, Dordrecht.
[13] Papadimitriou, C.H. and Steiglitz, K. (1982), *Combinatorial Optimization. Algorithms and Complexity*. Prentice Hall, Inc., Englewood Cliffs, N.J.
[14] Parrilo, P.A. (2000), *Structured Semidefinite Programs and Semi-algebraic Geometry Methods in Robustness and Optimization*. PhD thesis, California Institute of Technology, Pasadena, California, USA. Available at: http://www.cds.caltech.edu/~pablo/.
[15] Parrilo, P.A. and Sturmfels, B. (2001), *Minimizing polynomial functions*, Tech. Rep. math. OC/0103170, DIMACS, Rutgers University, March.

[16] Pólya, G. (1928), *Über positive Darstellung von Polynomen. Vierteljschr. Naturforsch. Ges. Zürich*, 73, 141–145 (also Collected Papers 2, 309–313, MIT Press, Cambridge, MA, London, 1974).
[17] Powers, V. and Reznick, B. (2001), A New Bound for Pólya's Theorem with Applications to Polynomials Positive on Polyhedra, *J. Pure Appl. Alg*. 164, 221–229.
[18] J. Renegar (2001), *A Mathematical View of Interior-Point Methods in Convex Optimization*. Forthcoming, SIAM, Philadelphia, PA.
[19] Quist, A.J., de Klerk, E., Roos, C. and Terlaky, T. (1998), Copositive Relaxation for General Quadratic Programming, *Optimization Methods and Software* 9, 185–209.
[20] Sturm, J.F. (1999), Using SeDuMi 1.02, a MATLAB Toolbox for Optimization Over Symmetric Cones, *Optimization Methods and Software* 11-12, 625–653.
[21] Vickers, G.T. and Cannings, C. (1988), Patterns of ESS's I, *J. Theor. Biol*. 132, 387–408.